

September 5, 2001

MEMORANDUM

SUBJECT: Multiple Use of Data Policy Statement

FROM: /s/ Margaret N. Schneider
Acting Chief Information Officer

TO: Quality Information Council (QIC) Members

As the Agency's Acting CIO, I am issuing the attached Policy Statement on Promoting Multiple Use of EPA Data through you to your respective program offices. Until the policy is established as a formal EPA Order, the policy statement should serve as an interim guideline for the management of our data collections.

You were briefed on the policy statement and the issues surrounding it at the September 2000 QIC meeting. At that time, Members were asked to provide feedback to the QIC Collections and Access Policy Subcommittee (QIC CAPS), who crafted the policy. This final draft incorporates the input received.

The next step will be to formalize the policy statement as an EPA Order. The issuance of the Order will ensure that:

- Agency funding which results in the collection or development of data also results in delivery of or access to the data
- All data will be managed as an asset
- Contract language and partnership agreements improve data availability and appropriate "share-ability"

The QIC CAPS will convene a workgroup to develop the EPA Order and create boilerplate language for contracts, IAGs, and cooperative agreements. At this time, I would like you to appoint a representative from your program offices to serve on the workgroup. So that the workgroup can move forward as soon as possible, **please send the names of your nominees to Suzanne Annand at Annand.Suzanne@EPA.gov, Mail Code 2841, or 202-260-6939, by September 21, 2001.**

I feel confident that EPA stands to gain significant benefits from adopting these basic principles and procedures. I look forward to working with you as we strive to optimize the management and use of the Agency's data collections.

Attachment

cc: QIC CAPS Members
Multi-Use Policy Subgroup Members

Policy Statement on Promoting Multiple Use of Environmental Data

Goal Statement

The Agency recognizes that large amounts of environmental data are collected and developed directly and indirectly through EPA-funded efforts. These efforts include data collected through contracts, grants, cooperative agreements,¹ interagency agreements, and direct field activities. Often, Agency program managers may not know of the existence of these data which could be used for their information needs and thus avoid costly duplication of data collections.

Particularly for projects in which EPA has substantial financial interest, the goal to ensure maximum sharing and use of such data can have a number of benefits including:

- efficient use of limited resources to collect data
- improved data quality
- faster dissemination of data and analyses thereof
- improved understanding of environmental issues
- elimination of potentially duplicative or overlapping data collection
- enhanced public participation and trust

The Quality and Information Council (QIC) recognizes that there are some immediate measures that could be taken to increase the multiple use of data collected by the Agency. These measures include development of basic principles and policies on use and sharing of data, a data collection clearinghouse function to review significant proposed data collection activities, implementation of short term measures such as standard contracts/grants language that would promote data sharing, and development of an EPA Order to provide a more detailed framework for achieving this goal. In order to achieve this goal, the QIC requests the Chief Information Officer (CIO) to issue the following policy statement.

Policy Statement

The CIO establishes the following basic principles on data sharing to be followed by all Agency program managers:*

¹Data collected under grants and cooperative agreements present unique challenges and opportunities. While EPA cannot award grants and cooperative agreements for the principal purpose of meeting the Agency's needs for data, EPA Programs can incidentally benefit from the use of this data beyond the purpose for which it was originally collected. Moreover, the Agency has broad rights to use, and authorize others to use, data collected under grants and cooperative agreements for Federal purposes.

1. Environmental data and information paid for by the Agency, or obtained under EPA supported projects:

- a. should be developed in accordance with the requirements of the mandatory EPA Quality Systems for environmental data (EPA Order 5360.1);
- b. should be accessible to the Agency and the public subject to Confidential Business Information, privacy and national security requirements;
- c. should be developed using adopted Agency data standards, formats and software. Where data standards are not available, or the use of EPA standards would be inconsistent with legal requirements, use appropriate industry practice standards or develop a documented interim data standard;
- d. should be developed with a perspective to integrating data and considering cross media environmental impacts;
- e. should provide metadata on new collections.

2. All EPA automated information application systems to be developed, which will potentially collect environmental data, are subject to the systems life cycle requirements as prescribed in Chapter 17 of the IRM Policy Manual, EPA Directive 2100.

3. Prior to embarking on new generations of any data or environmental information projects, the initiating office shall document in their project plans that it has consulted other sources or indexes of available data including, but not limited to, the Environmental Data Registry, the Paperwork Reduction Act Management System (PRAMS, the ICR Inventory), and the Environmental Information Management System (EIMS).

4. A system such as EIMS will act as the Agency mechanism for recording and inventorying data sources.

5. Metadata on past data collections should be collected and made available to the Agency and the public, on a selected basis, depending on priority (that is, utility of the data) and available resources.

** In some circumstances, EPA merely holds and/or stores data belonging to and voluntarily provided by others such as volunteer monitoring groups, research scientists, or other individuals. These provisions do not apply in these instances, but should be encouraged as good data management practices by everyone who collects data.*

Definitions

Environmental data are any measurements or information that describe environmental processes, location, or conditions, or ecological health effects and consequences, including information collected directly from measurements, produced from models, and compiled from other sources such as

data bases or the literature. Categories of this data include, but are not limited to, emissions and discharge data, ambient data (including biological inventory data),

public health and safety effects information, exposure data, geographical locational data, incident, and fish kill data.

Metadata is data about data. Librarians are very familiar with metadata development and application. Metadata can be thought of as a label on an object which provides descriptive information on the content of the object or as an index or template which describes electronic information on a server using data fields.

Implementation

An EPA Order shall be issued to make this policy mandatory. The Order will be developed with the cooperation of a cross-Agency workgroup. In addition to the Order, guidance to Agency staff will be provided regarding sources of information in existing collections. In the interim, this policy statement is effective immediately for all new and not yet approved ICRs, and for other projects not involving ICRs for which data collection has not yet begun. For those ICRs that have been approved, as well as for those information collections which are in progress, a phase-in approach will be used. Detailed implementation guidance will follow the issuance of the EPA Order. Included in this guidance will be several implementation activities summarized below:

- EPA offices are encouraged to consult PRAMS, EIMS, and other sources of information on current EPA data collections, in order to assess whether data they intend to collect have already been collected.
- EPA programs shall, to the extent that they can, review their existing collections and provide metadata about their collections to an existing system such as EIMS that OEI shall designate.
- Standard language (boilerplate) regarding data collection shall be developed that will be inserted into contracts, grants, IAGs, Cooperative Agreements and ITMRA submissions to ensure reasonable future use of the data the Agency collects.
- Conditions under which a waiver would be authorized for non-conformance with this policy shall be defined in the Order to be issued.

August 24, 2001